The Hecke Algebra $\mathbf{T}_k$ has Large Index

Frank Calegari
Matthew Emerton

February 1, 2008

Abstract

Let $\mathbf{T}_k(N)^{\text{new}}$ denote the Hecke algebra acting on newforms of weight k and level N. We prove that the power of p dividing the index of $\mathbf{T}_k(N)^{\text{new}}$ inside its normalisation grows at least linearly with k (for fixed N), answering a question of Serre. We also apply our method to give heuristic evidence towards recent conjectures of Buzzard and Mazur.

1 Introduction.

Let $\mathbf{F}$ be a finite field of characteristic p, and let $\overline{\rho} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathbf{F})$ be a modular Galois representation of tame level dividing N. Let $X(\overline{\rho})$ denote the space of global Galois deformations unramified outside primes dividing Np. (More precisely, by $X(\overline{\rho})$ we mean the rigid analytic generic fibre of the corresponding formal deformation space of $\overline{\rho}$.) A fundamental question first raised explicitly by Mazur is to understand the locus Ω of deformations inside $X(\overline{\rho})$ that are crystalline at p. Assuming the Fontaine–Mazur conjecture [9, Conj. 3c, p. 49], Ω is precisely the locus of modular points of level coprime to p. The theory of the Eigencurve [5] provides a nice deformation theory of such representations, but requires extra data: in particular, the U_p eigenvalue, or equivalently, a Frobenius eigenvalue of the corresponding Dieudonné module. The forgetful map from the eigencurve to $X(\overline{\rho})$ is at most two to one (see [13, Thm. 6.11]), but its image is certainly complicated, consisting as it does of an infinite union of modular arcs, each crossing to form the "infinite fern" of Mazur and Gouvêa [10]. A natural object of study is the topological closure $\overline{\Omega}$ of Ω (in the space of all $\overline{\mathbf{Q}}_p$-valued points of $X(\overline{\rho})$, with its usual p-adic topology). Two recent conjectures shed light on the structure of $\overline{\Omega}$.

For a deformation $\rho \in X(\overline{\rho})$, let $\mathbf{Q}_p(\rho)$ denote the field generated by the traces $\mathrm{Tr}(\rho(\mathrm{Frob}_\ell))$, for ℓ coprime to Np.

Conjecture 1.1 (Buzzard [2]) *There exists a constant c depending only on N and p such that for every $\rho \in \Omega$ (or equivalently, every $\rho \in \overline{\Omega}$),*

$$[\mathbf{Q}_p(\rho) : \mathbf{Q}_p] < c.$$

Conjecture 1.2 (Mazur) *Let $\rho \in \overline{\Omega}$ be a classical modular point. Then ρ is tamely potentially semistable at p. In other words, the local representation attached to ρ becomes semistable after restriction to a tame extension of $\mathbf{Q}_p$.*

Conjecture 1.1 rules out the possibility that $\overline{\Omega}$ arises as the points of some deformation ring R, since the rigid space associated to R would have to be at least one dimensional, and hence would contain points defined over arbitrarily large extensions of $\mathbf{Q}_p$. This should be contrasted with the fact that one does expect the set of crystalline representations whose Hodge–Tate weights lie in some fixed finite interval to be parameterised by a corresponding deformation space; indeed, assuming the Fontaine–Mazur conjecture, the resulting set is finite and the corresponding deformation space is given by a product of twisted Hecke algebras. Likewise, one should contrast Conjecture 1.2 with the results of L. Berger [1] and forthcoming work of Berger and P. Colmez, in which it is shown that the condition of being crystalline with bounded Hodge-Tate weights is closed in local deformation space. The subtlety of both conjectures thus lies in the fact that while the level N is fixed, the weights of the modular forms that define the set Ω are arbitrary.

This paper proves some results which provide theoretical evidence for Conjectures 1.1 and 1.2, and should also be of independent interest.

Acknowledgments: The authors would like to thank the anonymous referee for pointing out a stupidity in the proof of Lemma 3.3.

2 Results

Let R be a finite reduced $\mathbf{Z}_p$-algebra. We define the *index* of R to be the index $[\tilde{R} : R]$, where $\tilde{R}$ is the normalisation of R. We define the *discriminant* of R to be the discriminant of the trace pairing on R over $\mathbf{Z}_p$. We will apply these notions when R is one of various Hecke algebras.

For any level $N \geq 1$ and positive integer k, let $S_k(N)$ denote the space of cusp forms of weight k for the congruence subgroup $\Gamma_1(N)$ defined over $\overline{\mathbf{Q}}_p$, and let $S_k(N)^{\mathrm{new}}$ denote the subspace of S_k spanned by the newforms of conductor N. Let $\mathbf{T}_k(N)$ denote the finite $\mathbf{Z}_p$-algebra

$$\mathbf{T}_k(N) := \mathbf{Z}_p[T_2, \ldots, T_n, \ldots] \subseteq \mathrm{End}_{\mathbb{C}}(S_k(N))$$

generated by Hecke operators, and let $\mathbf{T}_k(N)^{\mathrm{new}}$ denote the quotient of $\mathbf{T}_k(N)$ that acts faithfully on $S_k(N)^{\mathrm{new}}$. The algebra $\mathbf{Q} \otimes \mathbf{T}_k(N)^{\mathrm{new}}$ is reduced, and has dimension $\dim_{\mathbb{C}} S_k(N)^{\mathrm{new}}$ over $\mathbf{Q}_p$. We denote the discriminant of $\mathbf{T}_k(N)^{\mathrm{new}}$ by $\Delta_k(N)^{\mathrm{new}}$, and its index by $I_k(N)^{\mathrm{new}}$. What happens to the index as the weight approaches infinity?

Theorem 2.1 *The p-adic valuation of the index grows at least linearly in k. Equivalently,*

$$\liminf_{k \to \infty} \frac{1}{k} \cdot \mathrm{ord}_p(I_k(N)^{\mathrm{new}}) \ > \ 0.$$

Let $\mathbf{T}_k(N)'$ denote the subring of $\mathbf{T}_k(N)$ generated by the Hecke operators T_n for n prime to N. This is a reduced subring of $\mathbf{T}_k(N)$, and so we may similarly consider its discriminant $\Delta_k(N)'$ and its index $I_k(N)'$. For any divisor M of N, the Hecke action on the space $S_k(M)^{\mathrm{new}}$ (thought of as a subspace of $S_k(N)$) induces a map $\mathbf{T}_k(N)' \to \mathbf{T}_k(M)^{\mathrm{new}}$, and the product map $\mathbf{T}_k(N)' \to \prod_{M|N} \mathbf{T}_k(M)^{\mathrm{new}}$ is injective, and becomes an isomorphism after tensoring with $\mathbf{Q}$, and hence after passing to normalisations. Thus $I_k(N)' \geq \prod_{M|N} I_k(M)^{\mathrm{new}}$, and so Theorem 2.1 implies that $\mathrm{ord}_p(I_k(N)')$ grows (at least) linearly in k. This answers positively a question that was first raised in print by Jochnowitz [12] (where she attributes the question to Serre). In that paper Jochnowitz also proves the weaker statement that $\mathrm{ord}_p(I_k(N)')$ becomes arbitrarily large as k approaches infinity. Since $\Delta_k(N)'$ (the discriminant of $\mathbf{T}_k(N)'$) is always divisible by $I_k(N)'$, Theorem 2.1 also has as a corollary that $\Delta_k(N)'$ grows (at least) linearly in k, a result which was already known from the work of Jochnowitz [12].

Our method of proving Theorem 2.1 is to first establish the following result:

Theorem 2.2 *Fix a level N, possibly divisible by p. Let f and g be two normalised cuspidal $\overline{\mathbf{Q}}_p$ eigenforms on $\Gamma_1(N)$ of arbitrary weight. Suppose moreover that f and g are residually congruent; that is, $f \equiv g \bmod \mathfrak{p}$, where $\mathfrak{p}$ denotes the maximal ideal of $\overline{\mathbf{Z}}_p$. Then there exists a rational number $\kappa > 0$ depending only on N and p such that for each n prime to p,*

$$a_n(f) \equiv a_n(g) \bmod p^{\kappa}\overline{\mathbf{Z}}_p.$$

The content of this theorem is the independence of κ from the weights of f and g. We view this result as some evidence towards Conjecture 1.1, since it shows that at least modulo p^{κ}, all eigenforms of fixed level have coefficients lying in some bounded extension of $\mathbf{Q}_p$.

The constant κ in Theorem 2.2 can be made explicit, and we use it to produce the following evidence towards Conjecture 1.2:

Theorem 2.3 *Assume the tame level $N = 1$. Let $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(F)$ be a modular point of $\overline{\Omega}$ of weight k. Assume moreover that $e(F) = 1$ (i.e. $F/\mathbf{Q}_p$ is unramified). Let $N(\rho)$ denote the conductor of ρ. If either*

1. *$p \geq 5$,*
2. *$p < 5$, $k = 2$, $N(\rho) \neq 27$,*

then ρ is tamely potentially semistable.

The case $p \geq 5$ of this theorem is an immediate consequence of Lemma 4.1 below (which when $F = \mathbf{Q}_p$ was proved as statement (d) on p. 67 of [9]). It is the cases when $p = 2$ or 3 that are more delicate, and which depend on certain explicit versions of Theorem 2.2.

3 Eigenforms modulo powers of p

The main difficultly in working with the space of eigenforms over $\overline{\mathbf{Z}}_p$ reduced modulo rational powers of p is that this space has no intrinsic geometric description. The Serre–Deligne lemma guarantees that residual eigenforms of weight at least 2 automatically lift to characteristic zero. This situation fails miserably, however, for eigenforms over Artin rings. For example, let f and g be two eigenforms in $S_k(\Gamma_1(N), \overline{\mathbf{Z}}_p)$ that are congruent modulo p but distinct modulo p^2. Then there exist infinitely many distinct eigenforms

$$(\alpha f + \beta g)/(\alpha + \beta) \in S_k(\Gamma_1(N), \overline{\mathbf{Z}}_p/p^2),$$

only finitely many of which can lift to eigenforms of weight k and characteristic zero. One way to avoid this problem is to not work directly with eigenforms at all, and to instead work with the entire space S_k of modular forms. Any congruences satisfied by Hecke operators on $S_k(\Gamma_1(N), \overline{\mathbf{Z}}_p/p^{\kappa})$ will then automatically be satisfied by eigenforms. This approach also has its difficulties, however.

For example, when $N = 1$ and $p = 2$, it can be shown using Coleman's theory [3] that elements of $S_k(\mathrm{SL}_2(\mathbf{Z}), \overline{\mathbf{Z}}_2/2)$ that are the reductions mod 2 of eigenforms in characteristic zero are always killed by T_2. Let us see how close we can get to this congruence using naive arguments: It is easy to show that the Hecke operator T_2 acts by zero on the space $S_k(\mathrm{SL}_2(\mathbf{Z}), \mathbf{F}_2)$. This implies in turn that T_2 induces a nilpotent operator on $S_k(\mathrm{SL}_2(\mathbf{Z}), \overline{\mathbf{Z}}_2/2)$. However, the operator T_2 on $S_k(\overline{\mathbf{Z}}_2/2)$ is highly non-semisimple, and the most that one can extract from this naive analysis is that $T_2^n = 0$, for some n that increases linearly with k. Correspondingly, one can infer from this only that for eigenforms f,

$$a_2(f) \equiv 0 \bmod 2^{1/n}\ \overline{\mathbf{Z}}_2,$$

a congruence which is not independent of the weight of f, and which is much weaker than the congruence cited at the beginning of the paragraph.

In order to overcome the difficulties of the type that occur in the naive argument of the preceding paragraph, we adopt an approach that is suggested by the arguments of Hatada [11]: Namely, we choose a better $\mathbf{T}_k(N)$-invariant lattice $\Lambda \subset S_k(N)$, for which the action of $\mathbf{T}_k(N)$ on $\mathbf{Z}_p/p \otimes \Lambda$ can be computed. Ideally, the action of $\mathbf{T}_k(N)$ on $\mathbf{Z}_p/p \otimes \Lambda$ will be semisimple, or as close to this as possible. It turns out that the space of modular symbols is well adapted to our purpose, and it is the method whereby Hatada [11] obtains congruences for modular forms of small level independent of the weight. By formalizing aspects of this argument in terms of geometry and cohomology, we prove Theorem 2.2.

Proof of Theorem 2.2. Without loss of generality we may assume that the congruence subgroup $\Gamma_1(N) \bigcap \Gamma(p)$ is torsion-free (by replacing N by an appropriate multiple if necessary). Let $Y(N,p)$ be the open modular curve that classifies elliptic curves with a fixed point of order N and full level-p structure. We impose no

conditions on the value of the Weil pairing on the basis elements giving the level-p structure, and so $Y(N,p)$ is typically a disconnected curve, with $\phi(p)$ connected components. Concretely, $Y(N,p)$ is equal to the quotient

$$\Gamma_1(N)\backslash(\mathcal{H}\times \mathrm{GL}_2(\mathbf{Z}/p)),$$

where $\mathcal{H}$ denotes the upper half-plane, and $\Gamma_1(N)$ acts on $\mathcal{H}$ through linear fractional transformations, and on $\mathrm{GL}_2(\mathbf{Z}/p)$ by left multiplication. Since $\Gamma_1(N)\bigcap\Gamma(p)$ is torsion-free, $Y(N,p)$ does in fact represent the appropriate moduli problem, and there exists a universal elliptic curve $\mathcal{E}$ over $Y(N,p)$. On $Y(N,p)$ we have the standard rank two local system $\mathcal{L}$, corresponding to the family of relative first cohomology groups of $\mathcal{E}$. Let $\mathcal{L}_k$ denote the kth symmetric power of $\mathcal{L}$; it is local system, free of rank $k+1$. If W_k denotes the kth symmetric power of the standard representation of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathbf{Z}^2$, then $\mathcal{L}_k$ has the following concrete description:

$$\mathcal{L}_k := \Gamma_1(N)\backslash(W_k\times\mathcal{H}\times \mathrm{GL}_2(\mathbf{Z}/p)).$$

Consider the short exact sequence of sheaves:

$$0\longrightarrow \mathcal{L}_k \xrightarrow{p\cdot} \mathcal{L}_k \longrightarrow \mathcal{L}_k/p\longrightarrow 0.$$

Taking cohomology, and remembering that $H^2(Y(N,p),\mathcal{L}_k)$ vanishes (since its dual $H^0_c(Y(N,p),\check{\mathcal{L}}_k)$ vanishes (here $\check{\mathcal{L}}_k$ denotes the $\mathbf{Z}$-dual of the local system of free $\mathbf{Z}$-modules $\mathcal{L}_k$), $Y(N,p)$ being an open Riemann surface), we obtain an isomorphism

$$H^1(Y(N,p),\mathcal{L}_k)/p\cong H^1(Y(N,p),\mathcal{L}_k/p).$$

This isomorphism is equivariant with respect to the prime-to-p Hecke operators.

The local system $\mathcal{L}_k/p$ (the reduction of $\mathcal{L}$ modulo p) is trivial over $Y(N,p)$, by definition of the moduli problem that $Y(N,p)$ represents. Thus

$$H^1(Y(N,p),\mathcal{L}_k/p)\cong W_k/p\otimes H^1(Y(N,p),\mathbf{Z}/p).$$

(The twisted coefficients $\mathcal{L}_k/p$ are actually untwisted, and so we can pull them out of the cohomology. The preceding concrete description of $\mathcal{L}_k$ shows that the $k+1$-dimensional $\mathbf{Z}/p$-vector space W_k/p is the fibre of $\mathcal{L}_k/p$ over any point of $Y(N,p)$.) Again, this isomorphism is equivariant for all the prime-to-p Hecke operators. (Where on the right hand side, these ignore the W_k/p factor, and just act on the second factor.) In particular, if n is coprime to p, and T_n is the nth Hecke operator acting on $H^1(Y(N,p),\mathcal{L}_k)/p$, then T_n satisfies (modulo p) a polynomial of degree independent of k, bounded explicitly by the dimension of $H^1(Y(N,p),\mathbf{Z}/p)$. The Eichler–Shimura isomorphism guarantees that any modular eigenform f of weight k and level N corresponds to an eigenform in $\mathbb{C}\otimes H^1(Y(N,p),\mathcal{L}_k)$. In particular, the eigenvalues of T_n acting on f will be among the eigenvalues of T_n acting on $H^1(Y(N,p),\mathcal{L}_k)$. Since $T_n \pmod p$ satisfies a fixed polynomial independent of k, we infer that there exists $\kappa>0$ such that the reduction modulo p^κ of any eigenvalue of T_n is determined explicitly by its residue in $\overline{\mathbf{F}}_p$. For example, one could take $1/\kappa$ to be the dimension of $H^1(Y(N,p),\mathbf{Z}/p)$. This proves Theorem 2.2. □

More precise values of κ can be extracted in particular cases. For example:

Lemma 3.1 *Assume that each connected component of $X(p)$ has genus 0. Let f be a cuspidal eigenform of weight k and level $\Gamma(p)$ (not necessarily a newform). Then for all primes $\ell \equiv \pm 1$ mod p,*

$$a_\ell(f) \equiv 1 + \ell \text{ mod } p\overline{\mathbf{Z}}_p.$$

Proof. If $X(p)$ has genus zero, then there is an isomorphism

$$H^1(Y(p), \mathbf{Z}/p) \simeq \widetilde{H}^0(C(p), \mathbf{Z}/p),$$

where the target of this isomorphism denotes the reduced 0-dimensional cohomology of the set of cusps $C(p) := X(p) \setminus Y(p)$. For $\ell \equiv \pm 1 \bmod p$, the action of T_ℓ on the cusps is given explicitly by $1 + \ell$. Thus T_ℓ satisfies the polynomial $T_\ell - 1 - \ell = 0$ on

$$W_k/p \otimes H^1(Y(p), \mathbf{Z}/p)$$

and the result follows. $\square$

Recall that the hypothesis of Lemma 3.1 is satisfied precisely for $p = 2, 3$, and 5. Lemma 3.1 (phrased as a statement for these values of p) was originally proved by Hatada [11], using a mixture of techniques. Our argument shows that Hatada's explicit computations ultimately rely on the fact that $X(p)$ has genus zero. However, our methods prove specific congruences for larger primes as well.

Lemma 3.2 *Let $p = 7$, and let f be a cuspidal eigenform of weight k and level $\Gamma(7)$ (not necessarily a newform). Then for all primes $\ell \equiv \pm 1$ mod 7,*

$$a_\ell(f) \equiv 1 + \ell \text{ mod } \sqrt{7} \cdot \overline{\mathbf{Z}}_7.$$

Proof. The curve $X(7)$ is the union of six connected components each of genus 3. There is an exact sequence:

$$0 \to H^1(X(7), \mathbf{Z}/7) \to H^1(Y(7), \mathbf{Z}/7) \to \widetilde{H}^0(C(7), \mathbf{Z}/7) \to 0.$$

The space $H^1(X(7), \mathbf{Z}/7)$ is "accounted for" by the cubic twists of the weight 2 form corresponding to the elliptic curve of conductor 49 and CM by $\mathbf{Z}[(1 + \sqrt{-7})/2]$. Explicitly one finds that $T_\ell - 1 - \ell$ is zero on $H^1(X(7), \mathbf{Z}/7)$ and on the cusps. Thus $(T_\ell - 1 - \ell)^2 = 0$ on

$$W_k/p \otimes H^1(Y(7), \mathbf{Z}/7)$$

and the result follows. $\square$

Serre conjectured [11] that the congruence would be satisfied with $\sqrt{7}$ replaced by 7, although we are not able to prove this.

By passing to other curves of non-trivial genus, and with a certain amount of non-trivial calculation we are able to prove other congruences, such as the following:

Lemma 3.3 *Let f be a cuspidal eigenform of weight k and level* 1*. Then for all primes $\ell \equiv \pm 1$ mod* 9*,*

$$a_\ell(f) \equiv 1 + \ell \ mod\ 9\overline{\mathbf{Z}}_3.$$

Proof. Let us consider to begin with an arbitrary integer $N \geq 3$; we will work on the modular curves $Y(N)$ and $Y(N^2)$ classifying elliptic curves with full level-N structure (respectively full level-N^2 structure). We use notation analogous to that introduced in the proof of Theorem 2.2.

Consider the commutative diagram

$$\begin{tikzcd} H^1_c(Y(N^2), \mathcal{L}_k/N^2) \arrow[r] \arrow[d] \arrow[dr] & H^1(Y(N^2), \mathcal{L}_k/N^2) \arrow[d] \\ H^1_c(Y(N), \mathcal{L}_k/N^2) \arrow[r] & H^1(Y(N), \mathcal{L}_k/N^2), \end{tikzcd} \tag{1}$$

in which the vertical arrows exist because cohomology (with or without compact supports) has a covariant functoriality for proper maps (such as the map $Y(N^2) \to Y(N)$), and in which the diagonal arrow is defined as the composite of top and right-hand (or equivalently, left-hand and bottom) arrows.

Claim. For any local system $\mathcal{F}$ of finite $\mathbf{Z}$-modules on $Y(N)$, the push-forward map $H^1(Y(N^2), \mathcal{F}) \to H^1(Y(N), \mathcal{F})$ is surjective. (Here we have used $\mathcal{F}$ also to denote the pull-back of $\mathcal{F}$ to $Y(N^2)$.)

Proof of claim. A simple dévisage, using the fact that $H^2(Y(N), \mathcal{F}) = 0$ (see the proof of Theorem 2.2), shows that it suffices to prove the claim in the case when $\mathcal{F}$ is a local system of $\mathbf{F}_\ell$-modules, for some prime ℓ. The map $H^1(Y(N^2), \mathcal{F}) \to H^1(Y(N), \mathcal{F})$ is then dual to the pull-back map $H^1_c(Y(N), \check{\mathcal{F}}) \to H^1_c(Y(N^2), \check{\mathcal{F}})$ (here $\mathcal{F}$ denotes $\mathbf{F}_\ell$-dual), and so it suffices to show that this map is injective. However, this latter map can be described in terms of modular symbols: namely, if $\mathrm{Div}^0(\mathbf{P}^1(\mathbf{Q}))$ denotes the group of divisors of degree zero supported on the elements of $\mathbf{P}^1(\mathbf{Q})$, then there is a commutative diagram

$$\begin{tikzcd} H^1_c(Y(N), \check{\mathcal{F}}) \arrow[r, "\sim"] \arrow[d] & \mathbf{F}_\ell[\pi_0(Y(N))]\check{} \otimes \mathrm{Hom}_{\Gamma(N)}(\mathrm{Div}^0(\mathbf{P}^1(\mathbf{Q})), \check{\mathcal{F}}) \arrow[d] \\ H^1(Y(N^2), \check{\mathcal{F}}) \arrow[r, "\sim"] & \mathbf{F}_\ell[\pi_0(Y(N^2))]\check{} \otimes \mathrm{Hom}_{\Gamma(N^2)}(\mathrm{Div}^0(\mathbf{P}^1(\mathbf{Q})), \check{\mathcal{F}}) \end{tikzcd}$$

in which the horizontal arrows are isomorphisms, and the right hand vertical arrow is the tensor product of the dual of the surjection $\mathbf{F}_\ell[\pi_0(Y(N^2))] \to \mathbf{F}_\ell[\pi_0(Y(N))]$ (induced by the map $Y(N^2) \to Y(N)$) and the inclusion $\mathrm{Hom}_{\Gamma(N)}(\mathrm{Div}^0(\mathbf{P}^1(\mathbf{Q})), \check{\mathcal{F}}) \subset \mathrm{Hom}_{\Gamma(N^2)}(\mathrm{Div}^0(\mathbf{P}^1(\mathbf{Q})), \check{\mathcal{F}})$. In particular, it is injective, and thus so is the left hand vertical arrow.

The preceding claim, applied to the local system $\mathcal{L}_k/N^2$, shows that the right hand vertical arrow of diagram 1 is surjective.

Let's make the following assumption:

Assumption. The diagonal arrow $H^1_c(Y(N^2), \mathcal{L}_k/N^2) \to H^1(Y(N), \mathcal{L}_k/N^2)$ of diagram 1 vanishes.

Claim. If the assumption holds, then for all primes $\ell \equiv \pm 1 \pmod{N^2}$, the operator $T_\ell - (1+\ell)$ annihilates $H^1(Y(N), \mathcal{L}_k/N^2)$.

Proof of claim. The local system $\mathcal{L}_k/N^2$ is trivial on $Y(N^2)$. Thus we may rewrite the map $H^1_c(Y(N^2), \mathcal{L}_k/N^2) \to H^1(Y(N^2), \mathcal{L}_k/N^2)$ as

$$W_k/N^2 \otimes H^1_c(Y(N^2), \mathbf{Z}/N^2) \to W_k/N^2 \otimes H^1(Y(N^2), \mathbf{Z}/N^2).$$

Knowing that $T_\ell - (1+\ell)$ kills the cusps of $X(N^2)$, we see that its image on $H^1(Y(N^2), \mathbf{Z}/N^2)$ lies in the image of $H^1_c(Y(N^2), \mathbf{Z}/N^2)$ in $H^1(Y(N^2), \mathbf{Z}/N^2)$. Thus the image of $T_\ell - (1+\ell)$ on $H^1(Y(N^2), \mathcal{L}_k/N^2)$ lies in the image of $H^1_c(Y(N^2), \mathcal{L}_k/N^2)$ in $H^1(Y(N^2), \mathcal{L}_k/N^2)$. Since the right hand vertical arrow of diagram 1 is surjective, we find that the image of $T_\ell - (1+\ell)$ on $H^1(Y(N), \mathcal{L}_k/N^2)$ is contained in the image of $H^1_c(Y(N^2), \mathcal{L}_k/N^2)$ by the diagonal arrow of that diagram. We are assuming that this vanishes, and so are done.

We turn to deriving sufficient conditions for the assumption to hold. To this end, we suppose that each connected component of $X(N)$ has genus zero, and consider the exact sequence of sheaves

$$0 \longrightarrow \mathcal{L}_k/N \xrightarrow{N\cdot} \mathcal{L}_k/N^2 \longrightarrow \mathcal{L}_k/N \longrightarrow 0.$$

Passing to cohomology, and also to cohomology with compact supports, on $Y(N)$, we obtain the following diagram with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & H^1_c(Y(N), \mathcal{L}_k/N) \arrow[r] \arrow[d] & H^1_c(Y(N), \mathcal{L}_k/N^2) \arrow[r] \arrow[d] & H^1_c(Y(N), \mathcal{L}_k/N) \arrow[d] & \\
& H^1(Y(N), \mathcal{L}_k/N) \arrow[r] & H^1(Y(N), \mathcal{L}_k/N^2) \arrow[r] & H^1(Y(N), \mathcal{L}_k/N) \arrow[r] & 0.
\end{tikzcd}$$

(As noted in the proof of Theorem 2.2, since $Y(N)$ is an open Riemann surface, H^0_c and H^2 with coefficients in any local system always vanish.)

Since $\mathcal{L}_k/N$ is trivial on $Y(N)$, we have isomorphisms

$$H^1_c(Y(N), \mathcal{L}_k/N) \cong W_k/N \otimes H^1_c(Y(N), \mathbf{Z}/N)$$

and

$$H^1(Y(N), \mathcal{L}_k/N) \cong W_k/N \otimes H^1(Y(N), \mathbf{Z}/N).$$

Since the components of $X(N)$ have genus zero, the natural map

$$H^1_c(Y(N), \mathbf{Z}/N) \to H^1(Y(N), \mathbf{Z}/N)$$

vanishes, and so in the preceding diagram the left-most and right-most vertical arrows both vanish. Thus the middle arrow of that diagram factors through the natural map

$$H^1_c(Y(N), \mathcal{L}_k/N^2) \to H^1_c(Y(N), \mathcal{L}_k/N),$$

and so the diagonal arrow of diagram 1 factors through the composite

$$H_c^1(Y(N^2), \mathcal{L}_k/N^2) \to H_c^1(Y(N), \mathcal{L}_k/N^2) \to H_c^1(Y(N), \mathcal{L}_k/N).$$

Again, using the fact that $\mathcal{L}_k/N^2$ is trivial on $Y(N^2)$, and that $\mathcal{L}_k/N$ is trivial on $Y(N)$, we may rewrite this composite as

$$W_k/N^2 \otimes H_c^1(Y(N^2), \mathbf{Z}/N^2) \to W_k/N \otimes H_c^1(Y(N), \mathbf{Z}/N).$$

This map is obtained by tensoring through the natural map $H_c^1(Y(N^2), \mathbf{Z}/N^2) \to H_c^1(Y(N), \mathbf{Z}/N)$ with the projection $W_k/N^2 \to W_k/N$. Thus to verify the assumption, it suffices to show that the former map vanishes.

Exploiting the isomorphism (for a Riemann surface) between H_c^1 and H_1, we may rewrite this map as a direct sum over the connected components of $X(N)$ of the maps

$$\Gamma(N^2)^{\mathrm{ab}}/N^2 \to \Gamma(N)^{\mathrm{ab}}/N.$$

Thus we see that our above assumption holds if the following two conditions are met: (i) Each component of $X(N)$ has genus zero; (ii) the image of $\Gamma(N^2)^{\mathrm{ab}}$ in $\Gamma(N)^{\mathrm{ab}}$ lies in $N\Gamma(N)^{\mathrm{ab}}$.

For any value of N, there is a natural isomorphism $\Gamma(N)/\Gamma(N^2) \cong M_2(\mathbf{Z}/N)^0$ (the additive group of traceless 2×2 matrices). Thus the commutator subgroup $\Gamma(N)^c$ of $\Gamma(N)$ is contained in $\Gamma(N^2)$, and so we have the short exact sequence

$$0 \to \Gamma(N^2)/\Gamma(N)^c \to \Gamma(N)^{\mathrm{ab}} \to M_2(\mathbf{Z}/N)^0 \to 0.$$

Claim. If N is such that $\Gamma(N)$ is free on three generators, then condition (ii) above holds.

Proof of claim. Since $\Gamma(N)$ is free on three generators, its abelianisation is a free $\mathbf{Z}$-module of rank three, and so the kernel of the surjection $\Gamma(N)^{\mathrm{ab}} \to M_2(\mathbf{Z}/N)^0$ must be precisely $N\Gamma(N)^{\mathrm{ab}}$, since $M_2(\mathbf{Z}/N)^0 \simeq (\mathbf{Z}/N\mathbf{Z})^3$.

If we note that each component of $X(3)$ is of genus zero, and that $\Gamma(3)$ is free on three generators, then we see that the preceding claim proves the lemma. □

This lemma allows one to prove (for example) that the eigenforms of weight 2 and level 243 attached to elliptic curves cannot be approximated by 3-adic eigenforms of level 1, since (in both cases) the coefficient a_{19} fails to satisfy the required congruence. As observed by Coleman and Stein [6], the congruences of Hatada can be used in a similar way to eliminate the possible 2-adic approximation by forms of level one of the unique form of weight 2 and level 32. In fact, by using Lemma 3.3 and Hatada [11], for $p = 2, 3$ one can show that all but one of the (finitely many) eigenforms of weight 2, level p^n (for $n \geq 3$) with coefficients lying in an unramified extension of $\mathbf{Q}_p$ cannot be p-adically approximated by eigenforms of level one. (The claim that there are finitely many such eigenforms is justified in the remark following Lemma 4.2.) The exception is the unique form of level 27. This last example does not seem especially anomalous;

it could be dealt with by a slight strengthening of Lemma 3.3. However, just as our method fails to establish the conjecture of Serre mentioned above, it also fails to prove the desired congruence. Thus, although we are able to prove congruences for all primes p and levels N, our method does not provide a "machine" for proving any particular congruence of this form.

Proof of Theorem 2.1. Let us fix the level N and the weight k, and also a maximal ideal $\mathfrak{m}$ of $\mathbf{T}_k(N)^{\mathrm{new}}$. We will write simply $\mathbf{T}$ to denote the localisation $\mathbf{T}_k(N)^{\mathrm{new}}_{\mathfrak{m}}$.

We write $\widetilde{\mathbf{T}}$ to denote the normalisation of $\mathbf{T}$; it is a product of a finite number, say d, of finite DVR extensions of $\mathbf{Z}_p$. Thus we may find an embedding of $\widetilde{\mathbf{T}}$ into $\overline{\mathbf{Z}}_p^d$ such that the image of $\mathbf{T}$ is contained in $\{(x_1, \dots, x_d) \in \overline{\mathbf{Z}}_p^d \,|\, x_1 \equiv \cdots \equiv x_d \bmod \mathfrak{p}\}$, where $\mathfrak{p}$ denotes the maximal ideal of $\overline{\mathbf{Z}}_p$. Let $\mathbf{T}^\circ$ denote the subring of $\mathbf{T}$ generated by the prime-to-p Hecke operators, let κ denote the rational number of Theorem 2.2, and let $p^\kappa \cap \widetilde{\mathbf{T}}$ denote the intersection of $\widetilde{\mathbf{T}}$ with the ideal in $\overline{\mathbf{Z}}_p^d$ generated by p^κ.

Claim: Let $\mathcal{O}$ denote the $\mathbf{Z}_p$-subalgebra of $\overline{\mathbf{Z}}_p$ obtained by projecting $\mathbf{T}^\circ$ onto the first factor of the product $\overline{\mathbf{Z}}_p^d$. If we regard $\mathcal{O}$ as being embedded diagonally into the product $\overline{\mathbf{Z}}_p^d$, then there is an inclusion $\mathbf{T}^\circ \subseteq \mathcal{O} + p^\kappa \cap \widetilde{\mathbf{T}}$.

Proof of claim: Each of the projections of $\mathbf{T}$ onto one of the factors of $\overline{\mathbf{Z}}_p^d$ determines a normalised Hecke eigenform, and by construction, these eigenforms are all congruent modulo $\mathfrak{p}$. The claim thus follows from Theorem 2.2.

The claim implies that

$$\mathbf{T} = \mathbf{T}^\circ[T_p] \subseteq \mathcal{O} + p^\kappa \cap \widetilde{\mathbf{T}} + \mathcal{O}[T_p],$$

while elementary algebra implies that $[\widetilde{\mathbf{T}} : \mathbf{T}]$ is divisible by $[\widetilde{\mathbf{T}}/(p^\kappa \cap \widetilde{\mathbf{T}}) : \mathbf{T}/(p^\kappa \cap \mathbf{T})]$. Yet

$$\mathbf{T}/(p^\kappa \cap \mathbf{T}) \subseteq \mathcal{O}/(p^\kappa \cap \mathcal{O}) + \mathcal{O}[T_p]/(p^\kappa \cap \mathcal{O}[T_p]).$$

The first term is manifestly finite, and Theorem 2.2 shows that it is independent of k. The theory of Coleman [3], [4] implies that the eigenforms of slope at most κ fit into finitely many analytic families. Thus the number of such forms is bounded independently of the weight, and consequently the second term is also finite and bounded independently of k.

On the other hand, if $\widetilde{\mathbf{T}}$ has rank n over $\mathbf{Z}_p$, then a simple argument shows that $\widetilde{\mathbf{T}}/(p^\kappa \cap \widetilde{\mathbf{T}})$ has order at least $p^{\kappa n}$, and thus $\mathrm{ord}_p[\widetilde{\mathbf{T}} : \mathbf{T}] > \kappa \cdot n - \epsilon$, for some explicitly computable constant ϵ, independent of k.

Now take the product over all maximal ideals of $\mathbf{T}_k(N)^{\mathrm{new}}$. From a well known result of Serre–Tate–Jochnowitz, the number of modular residual representations $\overline{\rho}$ unramified outside Np is finite, and this implies that the number of such maximal ideals is bounded independently of the weight k. On the other hand, the rank n of $\mathbf{T}_k(N)^{\mathrm{new}}$ over $\mathbf{Z}_p$ grows linearly with k, and so the same is true of the rank of its normalisation. Thus the index $I_k(N)^{\mathrm{new}}$ must grow at least linearly in k, and Theorem 2.1 is proved. $\square$

4 Crystalline Representations

We show in this section that there do not exist *any* wildly potentially semistable representations into $\mathrm{GL}_2(F)$ when $p \geq 5$, and F is absolutely unramified. Along with the discussion after Lemma 3.3 this suffices to prove Theorem 2.3. This result and the following lemma are undoubtedly known to the experts (indeed, the case of the lemma when $F = \mathbf{Q}_p$ and $d = 2$ is proved as statement (d) on p. 67 of [9]), but we include a proof for lack of a reference in the literature. We refer to [7] for a discussion of semistable p-adic representations of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ (and of the associated terminology and notation), and to [8] for a discussion of the Weil-Deligne group representations attached to them.

Lemma 4.1 *Let $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p) \rightarrow \mathrm{GL}_d(F)$ become semistable only after a wildly ramified extension. Let e denote the ramification index of F. Then $p \leq de + 1$.*

Proof. Denote the corresponding representation by V. Let $K/\mathbf{Q}_p$ be a minimal Galois extension such that the restriction of V to $\mathrm{Gal}(\overline{\mathbf{Q}}_p/K)$ is semistable. Let K_0 and F_0 denote the maximal unramified extensions of $\mathbf{Q}_p$ contained in K and F respectively. Without loss of generality, assume that $K_0 \subseteq F_0$ (clearly this does not affect $e(F)$). Let k_F be the residue field of K, and $\mathcal{O} = W(k_F)$ the ring of integers of F_0. Let

$$D := D_{st}(V|_K) = (B_{st} \otimes_{\mathbf{Q}_p} V)^{G_K}.$$

D is a free $K_0 \otimes_{\mathbf{Q}_p} F$ module of rank d. By assumption $K_0 \subseteq F$, and thus there is a natural isomorphism

$$K_0 \otimes_{\mathbf{Q}_p} F \simeq \bigoplus_{\mathrm{Hom}(K_0,F)} F.$$

If elements of $\mathrm{Hom}(K_0, F)$ are denoted by σ_i, this isomorphism is given explicitly by the map $a \otimes b \mapsto [a\sigma_i(b)]_i$. If $e_{\sigma_j} = [\delta_{ij}]_i$ is the jth idempotent, D naturally decomposes as a product

$$D \simeq \bigoplus_{\sigma_i} D_{\sigma_i}$$

where $D_{\sigma_i} := e_{\sigma_i} D$. Let $\sigma \in \mathrm{Hom}(K_0, F)$. There is a natural action of $\mathrm{Gal}(K/\mathbf{Q}_p)$ on D. This action is K_0 semi-linear, and does not preserve D_σ. Following [8], however, one may adjust this action (using the crystalline Frobenius) to obtain a *linear* representation of the Weil group $W_{\mathbf{Q}_p}$. In particular, the natural action of $\mathrm{Gal}(K/K_0)$ (equivalently, the inertia subgroup of $W_{\mathbf{Q}_p}$) on D is linear and preserves D_σ. Since D_σ is an F vector space of dimension d, we may also consider it as an F_0 vector space of dimension de. Choosing a suitable lattice inside D_σ for the action of $\mathrm{Gal}(K/K_0)$ we obtain a representation

$$\psi : \mathrm{Gal}(K/K_0) \rightarrow \mathrm{GL}_{de}(\mathcal{O}).$$

If ψ had non-trivial kernel, it would correspond to some Galois subfield $L \subset K$ such that $\mathrm{Gal}(K/L)$ acted trivially on D. Yet then by descent V would already be

semistable over L, contradicting the minimality assumption on K. In particular, there must exist an element of exact order p inside

$$\mathrm{GL}_{de}(\mathcal{O}).$$

The lemma then follows from the following well known (and easy) result. $\square$

Lemma 4.2 *Let $\mathcal{O}$ be a discrete valuation ring with maximal ideal generated by p. Suppose that $p > n+1$. Then $\mathrm{GL}_n(\mathcal{O})$ has no elements of exact order p.*

The preceding lemma is false if $p \leq n+1$. Consequently, if $p = 2$ or 3, it is possible to have potentially semistable representations $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p) \to \mathrm{GL}_2(\mathbf{Q}_p)$ that become semistable only after making a wildly ramified extension. However, a slightly more refined argument allows one to bound the conductor of such a representation, and hence to bound the power of p that divides the level of a newform whose Fourier coefficients lie in $\mathbf{Q}_p$. The required generalisation of Lemma 4.2 is that $\mathrm{GL}_n(\mathcal{O})$ contains (up to conjugacy) a finite number of finite p-groups. In particular, ρ becomes semistable after some explicitly computable extension, from which one can explicitly bound the conductor. In particular, for any fixed weight, there only finitely many newforms of p-power level with coefficients in $\mathbf{Q}_p$.

References

[1] L. Berger, *Limites de représentations cristallines*, Preprint.

[2] K. Buzzard, *Questions about slopes of modular forms*, To appear in Astérisque.

[3] R. Coleman, *Classical and overconvergent modular forms*, Invent. Math. **124** (1996) no. 1–3, 215–241.

[4] ———, *Classical and overconvergent modular forms*, Invent. Math. **127** (1997) no. 3, 417–479.

[5] ———, B. Mazur, *The eigencurve*, Galois representations in algebraic geometry (Durham, 1996), 1–113, London Math Soc. Lecture Note Ser. **254**, Cambridge Univ. Press, Cambridge, 1998.

[6] ———, W. Stein, *Approximation of Infinite Slope Modular Eigenforms By Finite Slope Eigenforms*, Dwork Proceedings (2002).

[7] J.-M. Fontaine, *Représentations p-adiques semi-stables*, Périodes p-adique, Astérisque **223**, Soc. Math. de France, Paris, 1994, 113–184

[8] ———, *Représentations l-adiques potentiallement semi-stables*, Périodes p-adique, Astérisque **223**, Soc. Math. de France, Paris, 1994, 321–347

[9] ———, B. Mazur, *Geometric Galois representations*, Conference on Elliptic curves and modular forms (Hong Kong, 1993), International Press, Cambridge, MA, 1995, 41–78

[10] F. Gouvêa, B. Mazur, *Families of modular eigenforms*, Math. Comp. **58** (1992), no. 198, 793–805.

[11] K. Hatada, *Eigenvalues of Hecke Operators on* $\mathrm{SL}(2,\mathbf{Z})$, Math. Ann. **239**, 1979, 75–96.

[12] N. Jochnowitz, *The Index of the Hecke Ring,* $\mathbf{T}_k$*, in the Ring of Integers of* $\mathbf{T}_k \otimes \mathbf{Q}$. Duke, **46**, 1979, no. 4, 861–869.

[13] M. Kisin, *Overconvergent modular forms and the Fontaine-Mazur conjecture*, Invent. Math. **153** (2003), 373–454.

Email addresses: `fcale@math.harvard.edu`
`emerton@math.northwestern.edu`